JOURNAL OF DIFFERENTIAL EQUATIONS 34, 496–522 (1979)

Spline Approximations for Functional Differential Equations

H. T. BANKS*

Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

AND

F. KAPPEL†

Institut für Mathematik, Universität Graz, Graz, Austria

Received October 3, 1978; revised February 12, 1979

We develop an approximation framework for linear hereditary systems which includes as special cases approximation schemes employing splines of arbitrary order. Numerical results for first- and third-order spline-based methods are presented and compared with results obtained using a previously developed scheme based on averaging ideas.

1. INTRODUCTION

In this paper we consider approximation techniques for functional differential equations (FDEs) based on classical "least squares" or best L_2 spline approximations. In a recent paper [7], Burns and Cliff developed an approximation scheme that employed piecewise linear function approximations. However, to our knowledge, the ideas for use of splines (piecewise linear or higher order) in approximating FDEs as developed in our presentation here are new. We develop our ideas in the context of an abstract approximation theorem (the Trotter–Kato approximation theorem in linear semigroup theory) which greatly facilitates arguments to establish convergence. Use of the Trotter–Kato theorem in this way is not new; indeed, it has been used recently in connection with other approximation schemes for problems involving linear and nonlinear FDEs (see [3] for a survey and, in addition, the recent papers [1, 2, 4, 7–9]).

* This research supported in part by the National Science Foundation under Grant NSF-GP-28931x3, and in part by the U.S. Air Force under contract AF-AFOSR-76-3092.

† This research supported by the National Science Foundation under grant NSF-GP-28931x3.

0022-0396/79/120496-27$02.00/0

We restrict our considerations in this paper to linear systems and their approximation. The use of the ideas presented here in nonlinear system problems along with applications to optimal control and parameter estimation problems will be discussed elsewhere. Our main purpose here is to develop the fundamental theoretical ideas for spline-based methods and present a sample of our findings in numerical experiments with these approximations. As will be evident from our discussions of the numerical results in Section 5 below, we believe that these spline-based approximation techniques can offer significant advantages over other methods (e.g., those discussed in [1, 3, 7–9]) in many instances.

We first develop in Section 2 a general setting for our system approximation problem in an appropriate Hilbert space. Much of the material in this section is closely related to known results, but must be presented in order to give a complete and clear discussion of our ideas. A general approximation result is developed in Section 3 and we then in Section 4 show how spline-based methods can be treated in a simple manner as a special case of the approximation ideas of Section 3.

One feature of our presentation is that the proofs in Section 4 follow immediately from standard results in spline theory. Aesthetically, our development has appeal since the theoretical foundations are cleaner than those for other types of methods proposed to date (e.g., see [3, 7]). From a practical viewpoint these methods, as we have already noted, should prove advantageous in many control and identification problems.

The following notation will be used thourghout the paper. For $-\infty < a < b < \infty$, $L_2(a, b; R^n)$ is the Hilbert space of equivalence classes of all functions $x: [a, b] \to R^n$ such that $|x|^2$ is integrable. $|\cdot|$ is the Euclidean norm in R^n, $\langle \cdot, \cdot \rangle_2$ and $|\cdot|_2$ denote the usual inner product and norm in $L_2(a, b; R^n)$. $C^k(a, b; R^n)$, $k = 0, 1, 2,...$, denotes the space of R^n-valued continuous functions which possess k continuous derivatives on $[a, b]$. (At end points of closed intervals, "derivative" of course means the appropriate one-sided derivative.) For $k = 0$ this is the usual space of continuous functions which we denote simply by $C(a, b; R^n)$. $W^{1,2}(a, b; R^n)$ is the space of absolutely continuous functions $x: [a, b] \to R^n$ such that $\dot{x} \in L_2(a, b; R^n)$. The space of equivalence classes of all functions $x: [0, \infty) \to R^n$ such that $|x|^2$ is integrable on bounded intervals will be denoted by $L_{2,\text{loc}} = L_{2,\text{loc}}(0, \infty; R^n)$. For any $t_1 > 0$ the collection of all $f \in L_{2,\text{loc}}$ restricted to $[0, t_1]$ gives the space $L_2(0, t_1; R^n)$, of course. In the special case where $a = -r$, $b = 0$ with $r > 0$ we shall abbreviate the notation above and simply write L_2, C^k, etc. The state space for our considerations will be the Hilbert space $Z = R^n \times L_2$ with norm $|(\eta, \varphi)|_Z = (|\eta|^2 + |\phi|^2)^{1/2}$, $(\eta, \phi) \in Z$, and inner product $\langle (\eta_1, \phi_1), (\eta_2, \phi_2) \rangle_Z = \eta_1{}^T \eta_2 + \langle \phi_1, \phi_2 \rangle_2$, $(\eta_i, \phi_i) \in Z$, $i = 1, 2$. Corresponding to C^k, $k = 0, 1, 2,...$, we introduce the linear subspace of Z defined by $\mathscr{C}^k = \{(\phi(0), \phi) \mid \phi \in C^k\}$. Again, for $k = 0$ we write $\mathscr{C}$ instead of $\mathscr{C}^0$. Throughout this paper we shall not distinguish notationally between a function ϕ and its equivalence class in L_2. If ϕ is

in C, then $\hat{\phi}$ denotes the element $(\phi(0), \phi)$ in $\mathscr{C}$. If on the other hand $\hat{\phi}$ is an element in $\mathscr{C}$, then ϕ denotes the unique element in C such that $\hat{\phi} = (\phi(0), \phi)$. The identity map in any space and its matrix representations will always be denoted by I.

For a function $x: [-r, \alpha) \to R^n$, $\alpha > 0$, the symbol x_t, $t \in [0, \alpha)$, denotes the function $[-r, 0] \to R^n$ defined by $x_t(\theta) = x(t + \theta)$, $\theta \in [-r, 0]$. Finally, for any function x of one independent variable we shall use interchangeably either $\dot{x}$ or Dx to denote the derivative of x with respect to this variable.

2. The Linear Functional Differential Equation

In this section we detail the type of equation we consider in this paper and state the results for this equation which are important for the developments in the following sections.

Given $(\eta, \phi) \in Z$ and $f \in L_{2,\text{loc}}$ we consider the Cauchy problem

$$\begin{aligned} \dot{x}(t) &= L(x_t) + f(t), \qquad t \geqslant 0, \\ (x(0), x_0) &= (\eta, \phi). \end{aligned} \tag{2.1}$$

For simplicity of exposition we assume that $L(x_t)$ has the form

$$L(x_t) = \sum_{i=0}^{m} A_i x(t - \tau_i) + \int_{-r}^{0} A(\theta)\, x(t + \theta)\, d\theta, \tag{2.2}$$

where $0 = \tau_0 < \cdots < \tau_m = r$ and A_i, $A(\theta)$ are $n \times n$ matrices, the elements of $\theta \to A(\theta)$ being square-integrable on $[-r, 0]$. Note that for any any right-hand side of type (2.2) we can assume $\tau_m = r$ (possibly after redefining $A(\theta)$ or with $A_m = 0$). The form of $L(x_t)$ given in (2.2) is, to our knowledge, sufficiently general to include all linear autonomous FDEs arising in applications. (In regard to more general equations see the remark at the end of this section.)

A solution of (2.1) is a function $x: [-r, \alpha) \to R^n$, $\alpha > 0$, which is absolutely continuous on $[0, \alpha)$ such that $(x(0), x_0) = (\eta, \phi)$ and for $0 \leqslant t < \alpha$,

$$x(t) = \eta + \int_0^t L(x_\sigma)\, d\sigma + \int_0^t f(\sigma)\, d\sigma. \tag{2.3}$$

A solution of (2.1) is denoted by $x(t) = x(t; \eta, \phi, f)$. In case $f \equiv 0$ we shall write simply $x(t; \eta, \phi)$.

Note that for any two functions $x, y: [-r, \alpha_1] \to R^n$, $\alpha_1 > 0$, which are absolutely continuous on $[0, \alpha_1]$ and satisfy $x_0 = y_0$ in L_2 and $x(t) = y(t)$ for $t \in [0, \alpha_1]$, the maps $\sigma \to L(x_\sigma)$ and $\sigma \to L(y_\sigma)$ are defined and are in the same equivalence class of R^n-valued functions on $[0, \alpha_1]$. In general L is not defined on all of Z, but can be defined on $\mathscr{C}$ by $L(\hat{\phi}) = L(\phi), \phi \in C$.

The following lemma is stated without its quite elementary proof.

LEMMA 2.1. *The solutions of* (2.1) *exist on* $[-r, \infty)$ *and are uniquely determined. Moreover, for sequences* $(\eta_k, \phi_k) \to (\eta, \phi)$ *in* Z *and* $f_k \to f$ *in* $L_{2,\mathrm{loc}}$ *we have for all* $T > 0$

$$\sup_{t\in[0,T]} |x(t) - x_k(t)| \to 0$$

as $k \to \infty$, *where* $x(t) = x(t; \eta, \phi, f)$ *and* $x_k(t) = x(t; \eta_k, \phi_k, f_k)$.

In the case of the homogeneous equation, i.e., $f \equiv 0$, we define the family $S(t)$, $t \geqslant 0$, of operators $Z \to Z$ by

$$S(t)(\eta, \phi) = (x(t; \eta, \phi), x_t(\eta, \phi)) \tag{2.4}$$

for each $(\eta, \phi) \in Z$.

LEMMA 2.2. (a) $S(t)$, $t \geqslant 0$, *is a* C_0*-semigroup of bounded linear operators.*

(b) *The infinitesimal generator* $\mathscr{A}$ *of this semigroup and its domain* $\mathscr{D}(\mathscr{A})$ *are given by*

$$\mathscr{D}(\mathscr{A}) = \{(\phi(0), \phi) \in Z \mid \phi \in W^{1,2}\},$$
$$\mathscr{A}(\phi(0), \phi) = L(\phi), \dot{\phi}), \qquad (\phi(0), \phi) \in \mathscr{D}(\mathscr{A}).$$

(c) *For* $k = 1, 2, \ldots$ *the sets*

$$\mathscr{D}^k = \{\hat{\phi} \in \mathscr{C}^k \mid \dot{\phi}(0) = L(\phi)\}$$

and

$$(\mathscr{A} - \lambda I)\,\mathscr{D}^k \qquad \textit{for } \lambda \textit{ sufficiently large}$$

are dense in Z.

Parts (a) and (b) of this lemma are by now standard results if one deals with FDEs in the state space Z (cf., for instance, [3, 9, 15]).

In order to establish density of $(\mathscr{A} - \lambda I)\,\mathscr{D}^k$ we first observe that $\mathscr{C}^k$, $k = 0, 1, 2, \ldots$, is dense in Z because $\mathscr{D}(\mathscr{A}^k) \subset \mathscr{C}^{k-1}$ and $\overline{\mathscr{D}(\mathscr{A}^k)} = Z$ for $k = 1, 2, \ldots$. This last fact is true for the infinitesimal generator of any C_0-semigroup of bounded linear operators in Z (see, for instance, any standard reference on linear semigroup theory such as [10] or [16]). Then density of $(\mathscr{A} - \lambda I)\,\mathscr{D}^k$ for λ sufficiently large follows once one argues

$$(\mathscr{A} - \lambda I)\,\mathscr{D}^k = \mathscr{C}^{k-1} \tag{2.5}$$

for $k = 1, 2, \ldots$ and such λ.

The resolvent operator $(\mathscr{A} - \lambda I)^{-1}$ exists for λ sufficiently large and is a bounded linear operator $Z \to \mathscr{D}(\mathscr{A})$. Therefore, given $\hat{\psi} \in \mathscr{C}^{k-1}$, the equation

$$(\mathscr{A} - \lambda I)\hat{\phi} = \hat{\psi} \tag{2.6}$$

has a unique solution $\hat{\phi}$ for λ sufficiently large. We only have to check if $\hat{\phi} \in \mathscr{D}^k$. But using the definition of $\mathscr{A}$ we see that (2.6) is equivalent to

$$\dot{\phi} - \lambda\phi = \psi \qquad \text{and} \qquad L(\phi) - \lambda\phi(0) = \psi(0),$$

which implies $\phi \in C^k$ and $\dot{\phi}(0) = \lambda\phi(0) + \psi(0) = L(\phi)$. This proves (2.5). Finally, density of $\mathscr{D}^k$ now follows at once from $\mathscr{D}^k = (\mathscr{A} - \lambda I)^{-1}\mathscr{C}^{k-1}$ for λ sufficiently large.

A fundamental notion we shall use in our approach is the concept of dissipativeness. Recall, that a linear operator $B: \mathscr{D}(B) \to Z$, $\mathscr{D}(B) \subset Z$, is called dissipative in Z if $\langle Bz, z\rangle_Z \leqslant 0$ for all $z \in \mathscr{D}(B)$. If B is the infinitesimal generator of a C_0-semigroup $T(t)$, $t \geqslant 0$, of bounded linear operators $Z \to Z$ and $B - \omega I$ is dissipative for some $\omega \in R$, then $| T(t)z |_Z \leqslant e^{\omega t} | z |_Z$ for all $z \in Z$ and $t \geqslant 0$. The solution semigroup $S(t)$, $t \geqslant 0$, defined in (2.4) in general satisfies an estimate

$$| S(t)z |_Z \leqslant Me^{\omega t} | z |_Z , \qquad z \in Z,\ t \geqslant 0,$$

where $\omega \in R$ and $M > 1$. Therefore $\mathscr{A} - \omega I$ will not be dissipative in Z for any $\omega \in R$. Following an idea used in [15] we introduce an equivalent norm in Z such that Z with this norm is again a Hilbert space Z_g and for some constant $\omega \in R$, $\mathscr{A} - \omega I$ is dissipative in the space Z_g. Roughly speaking, the new norm reflects the weight placed by (2.2) on different parts of the past history.

Corresponding to the difference part in (2.2) we define the weighting function g to be a step function on $[-r, 0]$ such that

$$g(\theta) = j \qquad \text{for} \quad \theta \in (-\tau_{m-j+1} , -\tau_{m-j}), \qquad j = 1,..., m.$$

The space Z supplied with the norm

$$|(\eta, \phi)|_{Z_g} = \left(| \eta |^2 + \int_{-r}^{0} | \phi(\theta)|^2 g(\theta)\, d\theta \right)^{1/2}, \qquad (\eta, \phi) \in Z,$$

will be denoted by Z_g. Obviously, Z_g is a Hilbert space with inner product

$$\langle(\eta_1 , \phi_1), (\eta_2 , \phi_2)\rangle_{Z_g} = \eta_1{}^T\eta_2 + \int_{-r}^{0} \phi_1(\theta)^T \phi_2(\theta)\, g(\theta)\, d\theta,$$

for $(\eta_i , \phi_i) \in Z$, $i = 1, 2$.

Following the arguments given in [3, p. 186], we can establish the following result.

LEMMA 2.3. *$\mathscr{A} - \omega I$ is dissipative in Z_g, i.e.,*

$$\langle \mathscr{A}z, z\rangle_{Z_g} \leqslant \omega | z |^2_{Z_g} \qquad \textit{for} \quad z \in \mathscr{D}(\mathscr{A}),$$

with

$$\omega = \frac{m+1}{2} + |A_0| + \frac{1}{2}\sum_{i=1}^{m} |A_i|^2 + \frac{1}{2}\int_{-r}^{0} |A(\theta)|^2 \, d\theta.$$

Obviously, the norm $|\cdot|_{Z_g}$ is equivalent to $|\cdot|_Z$. Therefore, all results stated in Z which involve only topological concepts remain valid in Z_g (e.g., the results of Lemma 2.2).

Lemma 2.3 will be basic for our use of the Trotter–Kato theorem in the next section in order to obtain an approximation result for the homogeneous Cauchy problem, i.e., (2.1) with $f \equiv 0$.

We now turn to the general nonhomogeneous case of (2.1) and define

$$z(t) \equiv z(t; \eta, \phi, f) \equiv S(t)(\eta, \phi) + \int_0^t S(t-\sigma)(f(\sigma), 0)\, d\sigma, \qquad t \geqslant 0 \quad (2.7)$$

for $(\eta, \phi) \in Z$ and $f \in L_{2,\mathrm{loc}}$. In applications (especially control problems) it is also important to investigate, for fixed $t_1 > 0$, the operator $\mathscr{F}: L_2(0, t_1; R^n) \to C(0, t_1; Z)$ defined for $0 \leqslant t \leqslant t_1$ and $f \in L_2(0, t_1; R^n)$ by

$$(\mathscr{F}f)(t) = \int_0^t S(t-\sigma)(f(\sigma), 0)\, d\sigma. \quad (2.8)$$

For our discussion below, the first part of the following lemma is an essential equivalence result. The second part, while not essential to our presentation here, is needed in applying our approximation results to control problems (see [2]), and is stated for future reference.

LEMMA 2.4. (a) *Given* $(\eta, \phi) \in Z$ *and* $f \in L_{2,\mathrm{loc}}$, *we have*

$$z(t) = (x(t; \eta, \phi, f), x_t(\eta, \phi, f)) \qquad \textit{for} \quad t \geqslant 0,$$

where $z(t)$ *is defined by* (2.7).

(b) *For any* $t_1 > 0$ *the operator* $\mathscr{F}$ *defined by* (2.8) *is a compact linear operator.*

The results of this lemma have already been established elsewhere (see Theorems 2.1 and 3.2 of [2], Theorem 1 of [4], Lemma I.3.7 of [11], or Lemma 2.2 of [12]) but for the sake of completeness we shall sketch the arguments needed to give a proof.

We first assume $(\phi(0), \phi) \in \mathscr{D}(\mathscr{A})$ and $f \in C^1(0, \infty; R^n)$. Then (see, for instance, [5, p. 31]) $z(t)$ given in (2.7) is the unique strong solution to the abstract Cauchy problem

$$\begin{aligned} \dot{z}(t) &= \mathscr{A}z(t) + (f(t), 0), \qquad t \geqslant 0, \\ z(0) &= (\phi(0), \phi). \end{aligned} \quad (2.9)$$

On the other hand for $x(t) = x(t; \phi(0), \phi, f)$ we have $x \in W^{1,2}(-r, t_1; R^n)$ for any $t_1 > 0$, which implies $x_t \in W^{1,2}$ for all $t \geqslant 0$. By a characterization of functions in $W^{1,2}$ (see [5, p. 21]), the derivative of the map $t \to x_t$ considered as a map $[0, \infty) \to L_2$ exists and $(d/dt)\, x_t = \dot{x}_t$, where $\dot{x}_t(\theta) = \dot{x}(t + \theta)$, $\theta \in [-r, 0]$. This together with (2.1) reveals that $(x(t), x_t)$ is also a solution of (2.9). By uniqueness of solutions of (2.9) we have $z(t) = (x(t), x_t)$, $t \geqslant 0$ and thus the equivalence in part (a) of Lemma 2.4 obtains whenever $(\eta, \phi) \in \mathscr{D}(\mathscr{A})$ and $f \in C^1(0, \infty; R^n)$.

To complete the proof of part (a) we observe that $\overline{\mathscr{D}(\mathscr{A})} = Z$ and $C^1(0, t_1; R^n)$ is dense in $L_2(0, t_1; R^n)$ for any $t_1 > 0$. Moreover, for $0 \leqslant t \leqslant t_1$, $z(t)$ given by (2.7) and $(x(t), x_t)$, $x(t) = x(t; \eta, \phi, f)$, depend continuously on $(\eta, \phi) \in Z$ and $f \in L_2(0, t_1; R^n)$.

In order to prove part (b) we fix $t_1 > 0$ and a bounded subset G of $L_2(0, t_1; R^n)$. We have to prove that, for each t in $[0, t_1]$, the set $\{(\mathscr{F}f)(t) \mid f \in G\}$ is precompact in Z and that $\{\mathscr{F}f \mid f \in G\}$ is bounded and equicontinuous.

If we observe that $(\mathscr{F}f)(t) = (x(t; 0, 0, f), x_t(0, 0, f))$, $t \in [0, t_1]$, and that $\{x(\cdot; 0, 0, f) \mid f \in G\}$ and $\{\dot{x}(\cdot; 0, 0, f) \mid f \in G\}$ are bounded subsets of $C(-r, t_1; R^n)$ and $L_2(-r, t_1; R^n)$, respectively, it is immediate that $\{(\mathscr{F}f)(t) \mid f \in G\}$ is precompact in $\mathscr{C}$ endowed with the norm $|\hat{\phi}|_{\mathscr{C}} = \sup_{-[J,0]} |\phi|$ and therefore also in Z. The same arguments also give equicontinuity and boundedness of $\{\mathscr{F}f \mid f \in G\}$.

Remark. All results given in this section remain valid for equations more general than those considered here. L must only be a continuous functional from C to R^n satisfying the conditions given by Borisovič and Turbabin in [6] (see also [2]). If $L(\phi) = \int_{-r}^{0} [d\eta(\theta)]\, \phi(\theta)$, $\phi \in C$, then the weighting function g in this more general case has to be defined by $g(\theta) = 1 + \int_{-r}^{\theta} |\, d\eta\, |$, $\theta \in [-r, 0]$.

3. A General Approximation Scheme

Basic to our approach is the following version of the so-called Trotter–Kato theorem (see [10, Theorem 4.6]).

Lemma 3.1. *Let $T(t)$ and $T^N(t)$, $N = 1, 2, \ldots$, $t \geqslant 0$ be C_0-semigroups in a Banach space Y with infinitesimal generators $\mathscr{B}$ and $\mathscr{B}^N$, respectively. Assume that the following conditions are satisfied.*

(i) (*Stability hypothesis*). *There exist constants ω and ω_N such that $\mathscr{B} - \omega I$ and $\mathscr{B}^N - \omega_N I$ are dissipative on Y and the sequence $\{\omega_N\}$ is bounded.*

(ii) (*Consistency hypothesis*). *There exists a subset $\mathscr{D} \subset \mathscr{D}(\mathscr{B}) \cap \bigcap_{N=1}^{\infty} \mathscr{D}(\mathscr{B}^N)$ which together with $(\mathscr{B} - \lambda I)\mathscr{D}$ for some $\lambda > 0$ is dense in Y and such that $\mathscr{B}^N y \to \mathscr{B} y$ for all $y \in \mathscr{D}$.*

Then

$$\lim_{N\to\infty} T^N(t)y = T(t)y$$

for all $y \in Y$ *uniformly on bounded t-intervals.*

We call $\{Z^N, P^N, \mathscr{A}^N\}$, $N = 1, 2,...$, an *approximation scheme* for the Cauchy problem (2.1) if $\{Z^N\}$ is a sequence of subspaces of Z_g, $\{P^N\}$ is the sequence of orthogonal projections $P^N: Z_g \to Z^N$, and $\{\mathscr{A}^N\}$ is a sequence of operators $Z \to Z^N$.

THEOREM 3.1. *Suppose that* $\{Z^N, P^N, \mathscr{A}^N\}$ *is an approximation scheme for* (2.1) *satisfying the following conditions*:

(i) $Z^N \subset \mathscr{D}(\mathscr{A})$, $N = 1, 2,...$.

(ii) $\mathscr{A}^N = P^N\mathscr{A}P^N$, $N = 1, 2,...$.

(iii) (a) $\lim_{N\to\infty} P^N z = z$ *in* Z *for all* $z \in Z$,

(b) *For some integer* $k \geqslant 1$ *we have* $\lim_{N\to\infty} L(\psi^N) = L(\psi)$ *in* R^n *and* $\lim_{N\to\infty} D\psi^N = D\psi$ *in* L_2 *for all* $\psi \in C^k$, *where* ψ^N *is defined by* $P^N\hat{\psi} = (\psi^N(0), \psi^N)$.

Then each $\mathscr{A}^N$ *is the infinitesimal generator of a* C_0*-semigroup* $S^N(t)$, $t \geqslant 0$, *such that*

$$S^N(t)\, Z^N \subset Z^N, \qquad N = 1, 2,...,$$

and

$$\lim_{N\to\infty} S^N(t)z = S(t)z \tag{3.1}$$

for all $z \in Z$, *uniformly on bounded t-intervals.*

Proof. We first show that $\mathscr{A}^N - \omega I$, $N = 1, 2,...$, is dissipative in Z_g with ω as given in Lemma 2.3. Since P^N and its dual map coincide, we get, by using the definition of $\mathscr{A}^N$ and $Z^N \subset \mathscr{D}(\mathscr{A})$, the estimate

$$\begin{aligned}\langle \mathscr{A}^N z, z\rangle_{Z_g} &= \langle \mathscr{A}P^N z, P^N z\rangle_{Z_g} \\ &\leqslant \omega \mid P^N z \mid^2_{Z_g} \leqslant \omega \mid z \mid^2_{Z_g}\end{aligned}$$

for all $z \in Z$, which proves the claim. We next observe that $\mathscr{A}P^N$ is closed and defined on all of Z and thus by the closed graph theorem [16] is bounded. It follows that $\mathscr{A}^N = P^N\mathscr{A}P^N$ is bounded and is the infinitesimal generator of a C_0-semigroup $S^N(t) = e^{\mathscr{A}^N t}$, $t \geqslant 0$. Invariance of Z^N under $S^N(t)$ is a consequence of $\mathscr{A}^N Z^N \subset Z^N$. In order to establish (3.1) it only remains to verify that hypotheses (i) and (ii) of Lemma 3.1 hold. Condition (i) follows from our comments above. In considering condition (ii), we choose $\mathscr{D} = \mathscr{D}^k$. By part (c)

of Lemma 2.2, $\mathscr{D}^k$ and $(\mathscr{A} - \lambda I)\mathscr{D}^k$, for λ sufficiently large, are dense in Z and therefore also in Z_g. For $\hat{\psi} \in \mathscr{D}^k$ we have the estimate

$$\begin{aligned} |\mathscr{A}^N\hat{\psi} - \mathscr{A}\hat{\psi}|_{Z_g} &\leqslant |P^N\mathscr{A}P^N\hat{\psi} - P^N\mathscr{A}\hat{\psi}|_{Z_g} + |P^N\mathscr{A}\hat{\psi} - \mathscr{A}\hat{\psi}|_{Z_g} \\ &\leqslant |\mathscr{A}P^N\hat{\psi} - \mathscr{A}\hat{\psi}|_{Z_g} + |P^N\mathscr{A}\hat{\psi} - \mathscr{A}\hat{\psi}|_{Z_g}. \end{aligned} \tag{3.2}$$

The second term on the right-hand side of (3.2) approaches zero as $N \to \infty$ by condition (iiia) of our hypothesis above. If we write $P^N\hat{\psi} = (\psi^N(0), \psi^N)$ then the first term on the right-hand side of (3.2) is $|(L(\psi^N), D\psi^N) - (L(\psi), D\psi)|_{Z_g}$, which tends to zero as $N \to \infty$ by (iiib). Note that convergence in Z is equivalent to convergence in Z_g. Therefore, we have $\lim_{N\to\infty} \mathscr{A}^N\hat{\psi} = \mathscr{A}\hat{\psi}$ for all $\hat{\psi} \in \mathscr{D}^k$, i.e., hypothesis (ii) of Lemma 3.1 is satisfied.

Remark 3.1. In order to prove $L(\psi^N) \to L(\psi)$ it suffices to show $\psi^N \to \psi$ in C (with sup-norm), because L is a continuous functional $C \to R^n$.

Remark 3.2. Since $\mathscr{C}^k$, $k = 0, 1, 2, \ldots$, is dense in Z, it is clear that (iiia) holds, if $\lim_{N\to\infty} P^N\hat{\psi} = \hat{\psi}$ for all $\hat{\psi} \in \mathscr{C}^k$ for some integer $k \geqslant 0$.

Remark 3.3. The proof of the Trotter–Kato theorem also yields estimates for the rate of convergence of $S^N(t)z \to S(t)z$, at least for certain $z \in Z$. This can be seen as follows. Inequalities (4.2) and (4.3) in [10, p. 88] imply

$$\begin{aligned} &|[S^N(t) - S(t)]\, R(\lambda_0 ; \mathscr{A})^2 z|_Z \\ &\quad\leqslant Me^{\omega T}\, |[R(\lambda_0 ; \mathscr{A}^N) - R(\lambda_0 ; \mathscr{A})]\, R(\lambda_0 ; \mathscr{A})z|_Z \\ &\qquad + Me^{\omega T} \int_0^T |[R(\lambda_0 ; \mathscr{A}^N) - R(\lambda_0 ; \mathscr{A})]\, S(\sigma)z|_Z\, d\sigma \\ &\qquad + |[R(\lambda_0 ; \mathscr{A}^N) - R(\lambda_0 ; \mathscr{A})]\, S(t)\, R(\lambda_0 ; \mathscr{A})z|_Z \end{aligned}$$

for $t \in [0, T]$, $T > 0$, $z \in Z$. Here λ_0 is a fixed real number such that $\lambda_0 > \omega$ and $R(\lambda; \mathscr{A})$, $R(\lambda; \mathscr{A}^N)$ is the usual notation for $(\lambda I - \mathscr{A})^{-1}$ and $(\lambda I - \mathscr{A}^N)^{-1}$, respectively. We also make use of the estimates

$$|S(t)| \leqslant Me^{\omega t}, \qquad |S^N(t)| \leqslant Me^{\omega t}, \qquad t \geqslant 0,$$

and $|R(\lambda_0; \mathscr{A}^N)| \leqslant M/(\lambda_0 - \omega)$ with some constant $M \geqslant 1$. Equation (4.14) in [10, p. 91] gives the estimate

$$|[R(\lambda_0 ; \mathscr{A}^N) - R(\lambda_0 ; \mathscr{A})]z|_Z \leqslant \frac{M}{\lambda_0 - \omega}\, |[\mathscr{A}^N - \mathscr{A}]\, R(\lambda_0 ; \mathscr{A})z|_Z$$

for $z \in Z$. Thus, for some constant $\tilde{M} = \tilde{M}(T)$ we finally obtain

$$\begin{aligned}
&|[S^N(t) - S(t)]\, R(\lambda_0 ; \mathscr{A})^2 z\,|_Z \\
&\quad \leqslant \tilde{M} \Big\{ |[\mathscr{A}^N - \mathscr{A}]\, R(\lambda_0 ; \mathscr{A})^2 z\,|_Z \\
&\qquad + \int_0^T |[\mathscr{A}^N - \mathscr{A}]\, S(\sigma)\, R(\lambda_0 ; \mathscr{A}) z\,|_Z \, d\sigma \\
&\qquad + |[\mathscr{A}^N - \mathscr{A}]\, S(t)\, R(\lambda_0 ; \mathscr{A})^2 z\,|_Z \Big\}
\end{aligned}$$

for all $t \in [0, T]$ and $z \in Z$. In order to use this estimate suppose that $\mathscr{M}$ is a subset of $\mathscr{D}(\mathscr{A}^2)$ such that there exists ρ so that for each $y \in \mathscr{M}$, there exists $\nu = \nu(y)$ with

$$|[\mathscr{A}^N - \mathscr{A}]\, y\,|_Z \leqslant \frac{\nu}{N^\rho}, \qquad N = 1, 2, \ldots.$$

Suppose $\mathscr{M}_1 \subset \mathscr{M}$ is such that $z \in \mathscr{M}_1$ implies

$$S(t) z \in \mathscr{M} \qquad \text{for all} \quad t \in [0, T],$$

and

$$S(t)(\lambda_0 I - \mathscr{A}) z \in \mathscr{M} \qquad \text{for all} \quad t \in [0, T],$$

with the constants $\nu = \nu(S(t)z)$, $\nu = \nu(S(t)(\lambda_0 I - \mathscr{A})z)$ being uniform in t. Then for each $z \in \mathscr{M}_1$ there exists a constant $\tilde{\nu} = \tilde{\nu}(z)$ such that

$$|[S^N(t) - S(t)]z\,|_Z \leqslant \frac{\tilde{\nu}}{N^\rho}, \qquad N = 1, 2, \ldots,$$

for all $t \in [0, T]$.

In the case of spline approximations (which are considered in Section 4), it is not difficult to find subsets $\mathscr{M}$ and $\mathscr{M}_1$ such that all the assumptions given above hold and thus convergence rate estimates as mentioned above are obtained. For example, define for $k \geqslant 1$ the sets

$$\mathscr{S}^k = \mathscr{D}(\mathscr{S}^k).$$

Then for the first-order splines discussed in Section 4 we may choose $\mathscr{M} = \mathscr{S}^2$ and $\mathscr{M}_1 = \mathscr{S}^3$ while for the cubic splines discussed there we may take $\mathscr{M} = \mathscr{S}^4$ and $\mathscr{M}_1 = \mathscr{S}^5$. However, the numerical results strongly indicate that the estimates one gets this way are not sharp at all. Certainly this problem needs further investigation.

We next turn to the nonhomogeneous problem and define for $(\eta, \phi) \in Z$ and $f \in L_{2,\mathrm{loc}}$

$$z^N(t; \eta, \phi, f) = S^N(t)\, P^N(\eta, \phi) + \int_0^t S^N(t - \sigma)\, P^N(f(\sigma), 0)\, d\sigma, \qquad (3.3)$$

$t \geqslant 0$, $N = 1, 2, \ldots$.

THEOREM 3.2. (a) *For* $(\eta, \phi) \in Z$ *and* $f \in L_{2,\mathrm{loc}}$ *we have*

$$\lim_{N\to\infty} z^N(t; \eta, \phi, f) = z(t; \eta, \phi, f), \qquad t \geqslant 0, \tag{3.4}$$

uniformly for t *in bounded intervals.*

(b) *For any* $t_1 > 0$ *the limit in* (3.4) *is uniform with respect to* $t \in [0, t_1]$ *and* f *in bounded subsets of* $L_2(0, t_1; R^n)$.

(c) *If* $\{f^k\}$ *is a sequence in* $L_2(0, t_1; R^n)$ *converging weakly to* f, *then*

$$\lim_{N,k\to\infty} z^N(t; \eta, \phi, f^k) = z(t; \eta, \phi, f)$$

uniformly for $t \in [0, t_1]$.

Proof. Of course, part (a) is a consequence of part (b). From Theorem 3.1 and $| S^N(t)z |_{Z_g} \leqslant e^{\omega t} | z |_{Z_g}$ for $t \geqslant 0$, $z \in Z$, and $N = 1, 2, \ldots$ it is clear that

$$\lim_{N\to\infty} S^N(t) P^N(\eta, \phi) = S(t)(\eta, \phi)$$

uniformly for $t \in [0, t_1]$.

Following [2] define the operators $T^N(t)\colon R^n \to Z$ and $T(t)\colon R^n \to Z$ by $T^N(t)\xi = S^N(t) P^N(\xi, 0)$ and $T(t)\xi = S(t)(\xi, 0)$, $\xi \in R^n$, $t \geqslant 0$. Then it is shown in [2, Lemma 3.2] that, for any $t \geqslant 0$, $T^N(t) \to T(t)$ as $N \to \infty$ in the uniform operator norm and $\int_0^t | T^N(\sigma) - T(\sigma)|^2 \, d\sigma \to 0$ as $N \to \infty$. Thus for $t \in [0, t_1]$

$$\begin{aligned}
&\left| \int_0^t [S^N(t-\sigma) P^N(f(\sigma), 0) - S(t-\sigma)(f(\sigma), 0)] \, d\sigma \right|_Z \\
&\qquad \leqslant \int_0^t | T^N(t-\sigma) - T(t-\sigma)| \, |f(\sigma)| \, d\sigma \\
&\qquad \leqslant \left\{ \int_0^{t_1} | T^N(\sigma) - T(\sigma)|^2 \, d\sigma \right\}^{1/2} |f|_{L_2(0,t_1;R^n)} ,
\end{aligned}$$

which proves part (b) of the theorem.

Part (c) follows from the inequality

$$\begin{aligned}
&| z^N(t; \eta, \phi, f^k) - z(t; \eta, \phi, f)|_Z \\
&\qquad \leqslant | z^N(t; \eta, \phi, f^k) - z(t; \eta, \phi, f^k)|_Z + | z(t; \eta, \phi, f^k) - z(t; \eta, \phi, f)|_Z ,
\end{aligned}$$

using part (b) above and the fact that $\{f^k\}$ is in a bounded subset of $L_2(0, t_1; R^n)$, and (b) of Lemma 2.4.

Remark 3.4. Theorem 3.2, part (c), is of special importance for the application of our approximation results to optimal control problems. See the discussions in [2, 3].

Remark 3.5. If we write $z^N(t; \eta, \phi, f) = (x^N(t), y^N(t))$, $t \geqslant 0$, then (3.4) implies

$$\lim_{N\to\infty} x^N(t) = x(t; \eta, \phi, f)$$

uniformly for $t \in [0, t_1]$, $t_1 > 0$. This is clear by definition of the norm in Z.

In order to obtain algorithms which can be implemented on a computer we assume from now on that

$$\dim Z^N = k_N < \infty, \qquad N = 1, 2, \ldots .$$

Then $z^N(t; \eta, \phi, f)$ given by (3.3) is the unique solution of the ordinary differential equation

$$\begin{aligned} \dot{z}^N(t) &= \mathscr{A}^N z^N(t) + P^N(f(t), 0), \qquad t \geqslant 0, \\ z^N(0) &= P^N(\eta, \phi) \end{aligned} \tag{3.5}$$

in Z^N.

We fix a basis $\hat{\beta}_1{}^N, \ldots, \hat{\beta}^N_{k_N}$ for Z^N. Since $Z^N \subset \mathscr{D}(\mathscr{A})$ we have $\hat{\beta}_j{}^N = (\beta_j{}^N(0), \beta_j{}^N)$, $j = 1, \ldots, k_N$, with $\beta_j{}^N \in W^{1,2}$. Define the $n \times k_N$ matrix function β^N by

$$\beta^N = (\beta_1{}^N, \ldots, \beta^N_{k_N})$$

and

$$\hat{\beta}^N = (\beta^N(0), \beta^N).$$

Then any $z^N \in Z^N$ can be written as

$$z^N = \hat{\beta}^N \alpha^N = (\beta^N(0)\, \alpha^N, \beta^N \alpha^N),$$

where $\alpha^N = \operatorname{col}(\alpha_1{}^N, \ldots, \alpha^N_{k_N}) \in R^{k_N}$ is the coordinate vector of z^N with respect to the chosen basis. The matrix representation of $\mathscr{A}^N$ restricted to Z^N with respect to this basis is denoted by A^N, while $w^N(t)$ and $F^N(t) = \operatorname{col}(F_1{}^N(t), \ldots, F^N_{k_N}(t))$ are the coordinate vectors of the solution $z^N(t)$ of (3.5) and $P^N(f(t), 0)$, respectively. That is, $z^N(t) = \hat{\beta}^N w^N(t)$ and $P^N(f(t), 0) = \hat{\beta}^N F^N(t)$. Then system (3.5) is equivalent to the system

$$\begin{aligned} \dot{w}^N(t) &= A^N w^N(t) + F^N(t), \qquad t \geqslant 0, \\ w^N(0) &= w_0{}^N, \end{aligned} \tag{3.6}$$

in R^{k_N}, where $\hat{\beta}^N w_0{}^N = P^N(\eta, \phi)$. If $w^N(t; w_0{}^N, F^N)$ denotes the solution of (3.6), then by Theorem 3.2 we have

$$\lim_{N\to\infty} \hat{\beta}^N w^N(t; w_0{}^N, F^N) = (x(t; \eta, \phi, f), x_t(\eta, \phi, f))$$

uniformly on bounded t-intervals and uniformly with respect to f in bounded subsets of $L_2(0, t_1; R^n)$, $t_1 > 0$. According to Remark 3.5 we also have

$$\lim_{N\to\infty} \beta^N(0)\, w^N(t; w_0{}^N, F^N) = x(t; \eta, \phi, f)$$

uniformly in $t \in [0, t_1]$, $t_1 > 0$.

In order to solve system (3.6) on a computer, one must, of course, know how to compute $P^N(\eta, \phi)$ and the matrix A^N. We first show how to compute the coordinate vector of $P^N(\eta, \phi)$ for $(\eta, \phi) \in Z$. Since P^N is the orthogonal projection $Z_g \to Z^N$ the element $P^N(\eta, \phi)$ is uniquely determined by the orthogonality relationship (in Z_g),

$$\{P^N(\eta, \phi) - (\eta, \phi)\} \perp Z^N.$$

This is equivalent to

$$\langle \hat{\beta}^N \alpha^N - (\eta, \phi), \hat{\beta}^N \rangle_{Z_g} = 0$$

or

$$Q^N \alpha^N = h^N(\eta, \phi), \tag{3.7}$$

where

$$Q^N = \langle \hat{\beta}^N, \hat{\beta}^N \rangle_{Z_g} = \beta^N(0)^T \beta^N(0) + \int_{-r}^{0} \beta^N(\theta)^T \beta^N(\theta)\, g(\theta)\, d\theta$$

and

$$h^N(\eta, \phi) = \langle \hat{\beta}^N, (\eta, \phi) \rangle_{Z_g} = \beta^N(0)^T \eta + \int_{-r}^{0} \beta^N(\theta)^T \phi(\theta)\, g(\theta)\, d\theta.$$

Thus, we see that in order to get the coordinate vector α^N of $P^N(\eta, \phi)$ we have to solve (3.7). Note that $(Q^N)^{-1}$ must exist since $P^N(\eta, \phi)$ is uniquely determined by (η, ϕ).

With respect to $F^N(t)$ we have $h^N(f(t), 0) = \beta^N(0)^T f(t)$ and therefore

$$Q^N F^N(t) = \beta^N(0)^T f(t).$$

Finally, the matrix A^N is calculated in the following manner. For $\hat{\phi}^N = (\phi^N(0), \phi^N) \in Z^N$ define $\alpha^N \in R^{k_N}$ and $\gamma^N \in R^{k_N}$ by

$$\hat{\phi}^N = \hat{\beta}^N \alpha^N \qquad \text{and} \qquad \mathscr{A}^N \hat{\phi}^N = \hat{\beta}^N \gamma^N.$$

Now, $\mathscr{A}^N \hat{\phi}^N = P^N \mathscr{A} \hat{\phi}^N = P^N(L(\phi^N), D\phi^N)$. In light of the calculations given above we see that γ^N is the solution of

$$Q^N \gamma^N = h^N(L(\phi^N), D\phi^N).$$

But by linearity of the map $(\eta, \phi) \to h^N(\eta, \phi)$ and L we have that

$$\begin{aligned} h^N(L(\phi^N), D\phi^N) &= h^N(L(\beta^N \alpha^N), (D\beta^N)\, \alpha^N) \\ &= H^N \alpha^N, \end{aligned}$$

where

$$H^N = h^N(L(\beta^N), D\beta^N) = \beta^N(0)^T L(\beta^N) + \int_{-r}^{0} \beta^N(\theta)^T (D\beta^N)(\theta)\, g(\theta)\, d\theta. \tag{3.8}$$

Thus, we obtain

$$A^N = (Q^N)^{-1} H^N.$$

Of course, on a computer one never actually computes $(Q^N)^{-1}$ but rather solves

$$Q^N \gamma^N = H^N \alpha^N \tag{3.9}$$

directly in order to obtain $\gamma^N = A^N \alpha^N$.

4. Spline Approximations

In this section we show that the general scheme given in Section 3 can be realized by choosing the Z^N as certain subspaces of spline functions. In order to keep the presentation brief we give the details only for first-order splines.

Corresponding to the partition $t_j^N = -j(r/N)$, $j = 0,\ldots, N$, of $[-r, 0]$ we define $Z_1^N = \{(\phi(0), \phi) \in \mathscr{C} \mid \phi$ is a first-order spline function with knots at t_j^N, $j = 0,\ldots, N\}$. A first-order spline function ϕ on $[-r, 0]$ with knots at the $\{t_j^N\}$ is simply a continuous function $[-r, 0]$ which is linear on each subinterval $[t_j^N, t_{j-1}^N]$, $j = 1,\ldots, N$. Let P_1^N be the orthogonal projection $Z_g \to Z_1^N$ and $\mathscr{A}_1^N = P_1^N \mathscr{A} P_1^N$, $N = 1, 2,\ldots$.

Theorem 4.1. *The approximation scheme* $\{Z_1^N, P_1^N, \mathscr{A}_1^N\}$ *satisfies all conditions of Theorem* 3.1 *and* $\dim Z_1^N = n(N + 1)$.

Proof. Conditions (i) and (ii) of Theorem 3.1 are trivially satisfied. It is also clear that $\dim Z_1^N = n(N + 1)$. For the verification of condition (iii) we take $k = 2$. We fix $\hat{\psi} \in \mathscr{C}^2$ and put $\hat{\psi}^N = P_1^N \hat{\psi}$. Let then ψ_1^N denote the interpolating first-order spline function defined by

$$\psi_I^N(t_j^N) = \psi(t_j^N), \qquad j = 0,\ldots, N.$$

Using the well-known convergence properties of interpolating splines (cf., for instance, [14, Theorem 2.5]) and the fact that $|\hat{\psi}^N - \hat{\psi}|_{Z_g} = \min_{\hat{\phi} \in Z_1^N} |\hat{\phi} - \hat{\psi}|_{Z_g}$ we obtain immediately that

$$|\hat{\psi}^N - \hat{\psi}|_{Z_g} \leqslant |\hat{\psi}_I^N - \hat{\psi}|_{Z_g} \leqslant (m)^{1/2} |\psi_I^N - \psi|_2 \leqslant O\left(\frac{1}{N^2}\right), \tag{4.1}$$

as $N \to \infty$. Thus, condition (iiia) of Theorem 3.1 is satisfied (cf. also Remark 3.2). The theorem in [14] quoted above also provides the estimate

$$| D(\psi_I{}^N - \psi)|_2 \leqslant O\left(\frac{1}{N}\right).$$

Therefore

$$\begin{aligned} | D(\psi^N - \psi)|_2 &\leqslant | D(\psi_I{}^N - \psi)|_2 + | D(\psi_I{}^N - \psi^N)|_2 \\ &\leqslant O\left(\frac{1}{N}\right) + M \cdot \frac{N}{r} | \psi_I{}^N - \psi^N |_2 , \end{aligned} \tag{4.2}$$

where for the second term on the right-hand side we have used the Schmidt inequality for polynomials of degree one on each of the intervals $[t_j{}^N, t_{j-1}^N]$, $j = 1,..., N$ (cf. [14, Theorem 1.5]). We observe that the constant M in this estimate is not dependent on N. Using

$$| \psi_I{}^N - \psi^N |_2 \leqslant | \psi_I{}^N - \psi |_2 + | \psi^N - \psi |_2 \leqslant O\left(\frac{1}{N^2}\right)$$

(note that $| \phi |_2 \leqslant | \hat{\phi} |_{Z_g}$ for $\phi \in C$) along with (4.2), we obtain finally the estimate

$$| D(\psi^N - \psi)|_2 \leqslant O\left(\frac{1}{N}\right) \tag{4.3}$$

as $N \to \infty$. This proves the second requirement in condition (iiib) of Theorem 3.1 and it only remains to verify that $L(\psi^N) \to L(\psi)$ as $N \to \infty$.

For $\theta \in [-r, 0]$ we have

$$\psi^N(\theta) = \psi^N(0) + \int_0^\theta (D\psi^N)(\sigma)\, d\sigma,$$

and hence

$$\begin{aligned} | \psi^N(\theta) - \psi(\theta)| &\leqslant | \psi^N(0) - \psi(0)| + \int_\theta^0 | D(\psi^N - \psi)(\sigma)|\, d\sigma \\ &\leqslant | \psi^N(0) - \psi(0)| + | \theta |^{1/2} \left(\int_\theta^0 | D(\psi^N - \psi)(\sigma)|^2\, d\sigma\right)^{1/2} \\ &\leqslant | \psi^N(0) - \psi(0)| + r^{1/2} | D(\psi^N - \psi)|_2 . \end{aligned}$$

This estimate together with (4.1) and (4.3) implies

$$| \psi^N(\theta) - \psi(\theta)| \leqslant O\left(\frac{1}{N}\right)$$

as $N \to \infty$, uniformly for $\theta \in [-r, 0]$. Therefore, $| L(\psi^N) - L(\psi)| \leqslant O(1/N)$ (see Remark 3.1) and the proof of Theorem 4.1 is completed.

Remark 4.1. The estimates given above reveal that

$$|\mathscr{A}_1^N\hat{\psi} - \mathscr{A}\hat{\psi}|_Z \leqslant O\left(\frac{1}{N}\right)$$

as $N \to \infty$ for $\hat{\psi} \in \mathscr{C}^2$. On the basis of this estimate the rate of convergence for $S^N(t)$ indicated in Remark 3.3 is

$$|S^N(t)\hat{\psi} - S(t)\hat{\psi}|_Z \leqslant O\left(\frac{1}{N}\right)$$

as $N \to \infty$, for $\hat{\psi}$ in $\mathscr{S}^3$.

Adhering to the general outline given in Section 3 for representation of system (3.5) in vector-matrix form, we use the following coordinate representation for our numerical calculations with first-order splines reported in Section 5 below.

Let e_j^N, $j = 0,..., N$, denote the scalar first-order spline function on $[-r, 0]$ characterized by

$$e_j^N(t_i^N) = \delta_{ij}, \qquad i, j = 0,..., N,$$

δ_{ij} being the Kronecker symbol. Then the matrix β^N is given by

$$\beta^N = (e_0^N,..., e_N^N) \otimes I,$$

where $\otimes$ denotes the Kronecker product and I is the $n \times n$ identity matrix. An element z in Z_1^N with coordinate vector α^N can be written as

$$z = \beta^N\alpha^N = \sum_{j=0}^{N} (e_j^N(0), e_j^N)\, a_j^N,$$

where the vectors $a_j^N \in R^n$ are such that $\alpha^N = \text{col}(a_0^N,..., a_N^N)$. For the case where $m = 1$ in (2.2) (where we have $g(\theta) \equiv 1$), simple calculations yield

$$Q_1^N = \frac{r}{N}\begin{bmatrix} N/r + \frac{1}{3} & \frac{1}{6} & 0 & \cdots & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ 0 & \cdots & 0 & \frac{1}{6} & \frac{1}{3} \end{bmatrix} \otimes I$$

for $N = 2, 3,...$ and

$$H_1^N = H_{11}^N + H_{12}^N,$$

where

$$H_{11}^N = \begin{bmatrix} A_0 + D_0{}^N & D_1{}^N & \cdots & D_{N-1}^N & A_1 + D_N{}^N \\ 0 & \cdot & \cdot & \cdot & 0 \\ \vdots & & & & \vdots \\ 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix}$$

with

$$D_j{}^N \equiv \int_{-r}^{0} A(\theta)\, e_j{}^N(\theta)\, d\theta = \int_{\max\{t_{j+1}^N, -r\}}^{\min\{t_{j-1}^N, 0\}} A(\theta)\, e_j{}^N(\theta)\, d\theta, \qquad j = 0,\ldots, N,$$

and

$$H_{12}^N = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & \cdots & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & & \\ 0 & & & & 0 \\ & & & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \cdots & 0 & \frac{1}{2} & & -\frac{1}{2} \end{bmatrix} \otimes I,$$

for $N = 2, 3,\ldots$.

Finally $F^N(t)$ is given by

$$F^N(t) = (Q_1{}^N)^{-1}\beta^N(0)^T f(t) = (Q_1{}^N)^{-1}\,\mathrm{col}(f(t),\ldots, 0).$$

It is the happy circumstance that all the ideas detailed above for first-order splines carry over to splines of higher order. For instance, in the case of cubic splines one may take $k = 4$ in order to prove a theorem analogous to Theorem 4.1. Instead of Theorem 2.5 in [14] we have to use Theorem 6.7 in [14]. It can be seen that $\dim Z_3{}^N = n(N + 3)$. For the construction of a basis of $Z_3{}^N$ we take the cubic de Boor splines (see [14, p. 73]). The matrix $Q_3{}^N$ is a seven-band matrix, $H_3{}^N$ is a six-band matrix, which we shall not give here explicitly (but see [13] for the details). For cubic splines we obtain the estimate

$$|\, \mathscr{A}_3{}^N \hat{\psi} - \mathscr{A}\hat{\psi}\,|_Z \leqslant O\left(\frac{1}{N^3}\right)$$

as $N \to \infty$ for $\hat{\psi} \in \mathscr{C}^4$. Use of the estimates in Remark 3.3 will yield

$$|\, S^N(t)\hat{\psi} - S(t)\hat{\psi}\,| \leqslant O\left(\frac{1}{N^3}\right)$$

as $N \to \infty$ for $\hat{\psi}$ in $\mathscr{P}^5$.

5. Examples

In this section, we discuss some of the examples we used in computations in order to investigate typical features of the spline approximations. For comparison, we also give the results for the so-called averaging approximations which are discussed in [2, 3]. In the following, we write AV, S_1, and S_3 for the averaging, first-order, and cubic spline approximations, respectively. Examples 1, 2, and 5 for AV and S_1 were computed on the IBM 360/67 at Brown University, whereas examples 3, 4, and 5 for S_3 were run on the UNIVAC 1100 of the Rechenzentrum Graz. We are grateful to Dr. D. Reber and Mr. P. Potter for their assistance with the calculations. Our main interest in performing the calculations was to demonstrate that the algorithm presented in Section 4 is numerically feasible and to obtain information about rates of convergence and accuracy of approximations. Therefore, the programs we used were not optimized with respect to computing time, storage, etc. As subroutines for solving (3.7), (3.9), and (3.6) we used standard algorithms such as Gaussian elimination and fourth-order Runge-Kutta.

In the tables below we use the following notation. δ_{AV}^N is one (appropriately chosen for the comparison being made) of the differences $|(w_1{}^N)_j(t) - x_j(t)|$, $j = 1,\ldots, n$, where $w_1{}^N(t)$ is the first segment of dimension n in the solution vector $w^N(t) = \mathrm{col}(w_1{}^N(t),\ldots, w_{N+1}^N(t))$ of the approximating ODE (see, for instance, [3, p. 179]) in case of AV and $x(t)$ is the true solution. Similarly, $\delta_{S_1}^N$ or $\delta_{S_3}^N$ is the appropriate choice from the differences

$$|(\beta^N(0)\, w^N)_j(t; w_0{}^N, F^N) - x_j(t; \eta, \phi, f)|, \qquad j = 1,\ldots, n,$$

when errors for the S_1 or S_3 approximations are being given.

Example 1. In this example, we study the equation for a damped oscillator with delayed restoring force and constant external force,

$$\ddot{x}(t) + \dot{x}(t) + x(t-1) = 10,$$

with initial conditions

$$x(\theta) = \cos\theta,\ \dot{x}(\theta) = -\sin\theta, \qquad -1 \leqslant \theta \leqslant 0.$$

The solution on the interval $[0, 2]$ is given by

$$x(t) = \mu(t) \equiv -9 - \sin 1 + 10t + (10 + \tfrac{1}{2}\sin 1 - \tfrac{1}{2}\cos 1)\, e^{-t} + \tfrac{1}{2}(\sin 1 - \cos 1)\sin t + \tfrac{1}{2}(\sin 1 + \cos 1)\cos t$$

for $t \in [0, 1]$ and

$$x(t) = \mu(t) - 29 - 2\sin 1 + \cos 1 + (19 + \sin 1)(t - 1) - 5(t - 1)^2 + (59/2 + \tfrac{3}{2}\sin 1 - \cos 1)\, e^{-(t-1)} + (10 + \tfrac{1}{2}\sin 1 - \tfrac{1}{2}\cos 1)(t - 1)\, e^{-(t-1)} + \tfrac{1}{2}(\sin 1 - 1)\cos(t - 1) + \tfrac{1}{2}(1 - \cos 1)\sin(t - 1)$$

for $t \in [1, 2]$.

For this example, the calculations were carried out for AV and S_1. Tables I and II show the numerical results for $x(t)$ and $\dot{x}(t)$. As is the usual practice, this second-order equation was converted for computational purposes into a 2×2 system of first-order equations for the unknown functions $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$. The results clearly show that the convergence for AV is like $1/N^{1-\epsilon}$, with ϵ a small positive number, whereas for S_1 it is like $1/N^2$. Another typical feature exhibited by this example is that the relative error for AV is

TABLE I

t	$x(t)$	δ^8_{AV}	δ^{16}_{AV}	δ^{32}_{AV}	$\delta^8_{S_1}$	$\delta^{16}_{S_1}$	$\delta^{32}_{S_1}$
0.25	1.27048	0.00126	0.00064	0.00032	0.00564	0.00132	0.00033
0.5	1.99367	0.00377	0.00194	0.00098	0.00895	0.00226	0.00056
0.75	3.06148	0.00601	0.00300	0.00151	0.01170	0.00292	0.00073
1.0	4.39272	0.00951	0.00421	0.00189	0.01346	0.00337	0.00084
1.25	5.92593	0.02266	0.01089	0.00527	0.01478	0.00374	0.00094
1.5	7.60007	0.05099	0.02646	0.01357	0.01281	0.00307	0.00075
1.75	9.34402	0.09117	0.04813	0.02484	0.00763	0.00200	0.00050
2.0	11.08330	0.1375	0.0725	0.0373	0.0043	0.0010	0.00030

TABLE II

t	$\dot{x}(t)$	δ^8_{AV}	δ^{16}_{AV}	δ^{32}_{AV}	$\delta^8_{S_1}$	$\delta^{16}_{S_1}$	$\delta^{32}_{S_1}$
0.25	2.06969	0.00882	0.00453	0.00230	0.00721	0.00221	0.00050
0.5	3.64428	0.01006	0.00521	0.00267	0.00526	0.00120	0.00028
0.75	4.84445	0.00851	0.00319	0.00135	0.00156	0.00057	0.00014
1.0	5.76581	0.02570	0.01038	0.00389	0.00083	0.00011	0.00002
1.25	6.45956	0.08326	0.04532	0.02428	0.00054	0.00003	0.00007
1.5	6.88559	0.14066	0.07695	0.04057	0.00977	0.00319	0.00063
1.75	7.01599	0.17685	0.09403	0.04835	0.00853	0.00184	0.00045
2.0	6.84972	0.19072	0.09909	0.05029	0.00611	0.00160	0.00044

increasing much faster with time (in this example for AV the relative error for $t = 2$ is about seven times the relative error at $t = 0.25$) than it is for S_1.

EXAMPLE 2. Here we deal with the equation for the oscillator with delayed damping

$$\ddot{x}(t) + \dot{x}(t-1) + x(t) = 1$$

with initial data

$$x(\theta) = \dot{x}(\theta) = 0 \qquad \text{for} \quad \theta \in [-1, 0].$$

The solution on [0, 2] is

$$x(t) = 1 - \cos t \qquad \text{for} \quad t \in [0, 1]$$

and

$$x(t) = 1 - \cos t + \tfrac{1}{2}(t-1)\cos(t-1) - \tfrac{1}{2}\sin(t-1) \qquad \text{for} \quad t \in [1, 2].$$

Again, the calculations were done for AV and S_1. The data given in Tables III and IV depict behavior similar to that found in Example 1.

TABLE III

t	$x(t)$	δ^8_{AV}	δ^{16}_{AV}	δ^{32}_{AV}	$\delta^8_{S_1}$	$\delta^{16}_{S_1}$	$\delta^{32}_{S_1}$
0.25	0.031088	0	0	0	0.000661	0.000157	0.000040
0.5	0.122417	0.000016	0	0	0.001185	0.000298	0.000074
0.75	0.268311	0.000447	0.000085	0.000010	0.001609	0.000401	0.000100
1.0	0.459698	0.003599	0.001478	0.000580	0.001767	0.000444	0.000110
1.25	0.682090	0.012699	0.006737	0.003529	0.001985	0.000481	0.000122
1.5	0.908946	0.023928	0.012988	0.006817	0.001223	0.000293	0.000072
1.75	1.111810	0.033150	0.017880	0.009290	0.000260	0.000070	0.000020
2.0	1.265563	0.037213	0.02003	0.010383	0.001887	0.000477	0.000117

TABLE IV

t	$\dot{x}(t)$	δ^8_{AV}	δ^{16}_{AV}	δ^{32}_{AV}	$\delta^8_{S_1}$	$\delta^{16}_{S_1}$	$\delta^{32}_{S_1}$
0.25	0.247404	0.000001	0	0	0.001138	0.000314	0.000074
0.5	0.479426	0.000283	0.000017	0.000001	0.000913	0.000218	0.000054
0.75	0.681639	0.004566	0.001231	0.000214	0.000148	0.000081	0.000018
1.0	0.841471	0.024575	0.013283	0.006999	0.000642	0.000216	0.000078
1.25	0.918059	0.044008	0.025291	0.013826	0.001180	0.000122	0.000041
1.5	0.877639	0.043160	0.023308	0.011891	0.002876	0.000854	0.000202
1.75	0.728371	0.028576	0.014968	0.007548	0.004395	0.001004	0.000255
2.0	0.488562	0.001830	0.000834	0.000551	0.004057	0.001039	0.000265

EXAMPLE 3. We consider a two-dimensional system described by the equation

$$\dot{x}(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} x(t-1)$$

with initial condition

$$x(\theta) = \operatorname{col}(0, \sin 2\pi\theta) \qquad \text{for} \quad \theta \in [-1, 0].$$

The solution on [0, 2] is given by

$$x_1(t) = \frac{1}{(2\pi)^2 - 1}(-2\pi \sin t + \sin 2\pi t),$$
$$x_2(t) = \dot{x}_1(t)$$

for $t \in [0, 1]$ and

$$x_1(t) = \frac{2\pi}{(2\pi)^2 - 1}\left\{-\sin t + \sin(t-1) - \frac{1}{(2\pi)^2 - 1}\cos(t-1) + \frac{1}{2}(t-1)\sin(t-1) + \frac{1}{(2\pi)^2 - 1}\cos 2\pi t\right\},$$
$$x_2(t) = \dot{x}_1(t)$$

for $t \in [1, 2]$. In Tables V and VI we give the numerical results for AV and S_3. For AV we have convergence like $1/N^{1-\epsilon}$, $\epsilon > 0$ small. The data for cubic splines do not behave as regularly as the data for AV or as that for S_1 in the previous examples. A possible explanation for this is the more complicated structure of the matrices which one encounters in S_3. For instance, one must be very careful when dealing with Eq. (3.9). We note that $N = 4$ for S_3 gives results comparable to that for $N = 128$ for AV.

EXAMPLE 4. We next use an example (due to Popov) for a degenerate system where we have $(1, -2, -1)^T x(t; \eta, \phi) = 0$ for $t \geqslant 2$ and all $(\eta, \phi) \in Z$. The equation is

$$\dot{x}(t) = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} x(t-1).$$

We choose the discontinuous initial data

$$\eta = \operatorname{col}(1, 1, 1) \qquad \text{and} \qquad \phi(\theta) = 0 \qquad \text{for} \quad \theta \in [-1, 0).$$

TABLE V

t	$x_1(t)$	δ^8_{AV}	δ^{16}_{AV}	δ^{32}_{AV}	δ^{128}_{AV}	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$
0.2	−0.0077243	0.0031729	0.0019945	0.0011167	0.0003036	0.0001220	0.0001100	0.0000191
0.4	−0.0483129	0.0049332	0.0023202	0.0010570	0.0002322	0.0028794	0.0000757	0.0000069
0.6	−0.1074768	0.0263038	0.0158887	0.0088920	0.0024340	0.0022146	0.0000823	0.0000027
0.8	−0.1418545	0.0336931	0.0199680	0.0111845	0.0031488	0.0008733	0.0000200	0.0000014
1.0	−0.1374048	0.0194136	0.0083610	0.0031054	0.0002812	0.0007086	0.0000233	0.0000063
1.2	−0.1193564	0.0084475	0.0021558	0.0002458	0.0000775	0.0018660	0.0001640	0.0000233
1.4	−0.0919509	0.0024952	0.0008709	0.0006891	0.0003232	0.0027533	0.0001455	0.0000167
1.6	−0.0502958	0.0064887	0.0032927	0.0013573	0.0002243	0.0006033	0.0001682	0.0000010
1.8	0.0033271	0.0198480	0.0113461	0.0062237	0.0017455	0.0013331	0.0000795	0.0000129
2.0	0.0595777	0.0323715	0.0185864	0.0102059	0.0028011	0.0004707	0.0000320	0.0000108

TABLE VI

t	$x_2(t)$	δ^8_{AV}	δ^{16}_{AV}	δ^{32}_{AV}	δ^{128}_{AV}	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$
0.2	−0.1095765	0.0065100	0.0066771	0.0045198	0.0013945	0.0097339	0.0007247	0.0001672
0.4	−0.2825064	0.0899860	0.0538617	0.0294842	0.0078536	0.0091217	0.0007947	0.0003820
0.6	−0.2668753	0.0955451	0.0623764	0.0374918	0.0110575	0.0152023	0.0014783	0.0001167
0.8	−0.0633063	0.0303768	0.0294751	0.0213255	0.0067536	0.0042850	0.0010228	0.0001734
1.0	0.0750646	0.0802746	0.0916800	0.0372904	0.0118155	0.0089749	0.0019726	0.0004590
1.2	0.1085746	0.0347772	0.0107840	0.0006172	0.0023401	0.0054677	0.0026357	0.0002546
1.4	0.1705014	0.0323881	0.0092470	0.0012010	0.0004231	0.0004590	0.0007226	0.0003701
1.6	0.2441385	0.0585545	0.0331494	0.0194799	0.0059646	0.0143525	0.0009849	0.0004964
1.8	0.2833447	0.0697551	0.0425845	0.0254048	0.0077916	0.0029493	0.0018694	0.0001235
2.0	0.272566	0.0525860	0.0277064	0.0131444	0.0023532	0.0053228	0.0025587	0.0003502

The solution is given by

$$x(t) = \text{col}(1 + 2t - t^2,\ 1 - t,\ 1) \qquad \text{for} \quad t \in [0, 1],$$
$$x(t) = \text{col}(2,\ 0,\ -2 + 4t - t^2) \qquad \text{for} \quad t \in [1, 2],$$

and

$$x(t) = \text{col}(2, 0, 2) \qquad \text{for} \quad t \geqslant 2.$$

Here the initial data are in the subspaces Z^N for AV but not for S_3. Thus we might expect that AV will initially (i.e., for t small) give a better approximation than S_3. Due to the simple algebraic structure of the approximating ODEs in the case of AV it is not difficult to prove that we have

$$\lim_{t \to \infty} w_1{}^N(t) = \text{col}\left(2 + \frac{1}{N},\ 0,\ 2 + \frac{1}{N}\right)$$

for all $N = 1, 2, \ldots$. So for AV we also can expect a very good approximation of $x_2(t)$ for t large. The error for $x_1(t)$ and $x_3(t)$ should tend to $1/N$ as $t \to \infty$. That this is true is shown by the numerical data displayed in Tables VII, VIII, and IX. Again, a characteristic feature of S_3 is that the error is not increasing in magnitude over the time interval under consideration. Note that S_3 for $N = 4$ gives a better approximation of $x_1(t)$ for $t \geqslant 1$ and of $x_3(t)$ for $t \geqslant 2.6$ than AV for $N = 128$. With respect to $x_2(t)$ we can see that S_3 for $N = 4$ gives a better approximation than AV for $N = 128$ around $t = 1$, where the true solution $x_2(t)$ has a jump discontinuity in the derivative. That AV behaves better for

TABLE VII

t	$x_1(t)$	δ_{AV}^{32}	δ_{AV}^{64}	δ_{AV}^{128}	$\delta_{S_3}^{4}$	$\delta_{S_3}^{8}$	$\delta_{S_3}^{16}$
0.2	1.36000	0	0	0	0.01211	0.00493	0.00247
0.4	1.64000	0.00001	0	0	0.00885	0.00528	0.00001
0.6	1.84000	0.00001	0	0	0.00446	0.00334	0.00190
0.8	1.96000	0.00105	0.00015	0.00001	0.00980	0.00074	0.00146
1.0	2.0	0.01416	0.00729	0.00372	0.00020	0.00729	0.00372
1.2		0.02866	0.01519	0.00777	0.00725	0.00424	0.00175
1.4		0.03100	0.01562	0.00781	0.00160	0.00460	0.00130
1.6		0.03123	0.01563	0.00781	0.00527	0.00406	0.00085
1.8		0.03125	0.01563	0.00781	0.00331	0.00281	0.00242
2.0	⋮				0.00256	0.00117	0.00130
2.2					0.00306	0.00049	0.00096
2.4		⋮	⋮	⋮	0.00053	0.00178	0.00214
2.6					0.00212	0.00246	0.00115
2.8					0.00054	0.00251	0.00092
3.0	2.0	0.03125	0.01563	0.00781	0.00117	0.00200	0.00191

TABLE VIII

t	$x_2(t)$	δ_{AV}^{32}	δ_{AV}^{64}	δ_{AV}^{128}	$\delta_{S_3}^{4}$	$\delta_{S_3}^{8}$	$\delta_{S_3}^{16}$
0.2	0.8	0	0	0	0.00947	0.00536	0.00230
0.4	0.6	0.00001	0	0	0.01206	0.00452	0.00035
0.6	0.4	0.00017	0	0	0.00526	0.00260	0.00202
0.8	0.2	0.00881	0.00200	0.00020	0.00153	0.00015	0.00139
1.0	0	0.07034	0.04980	0.03524	0.02133	0.01895	0.00949
1.2	0	0.01375	0.00384	0.00056	0.01505	0.00168	0.00001
1.4		0.00159	0.00009	0	0.01369	0.00068	0.00094
1.6		0.00012	0	0	0.00105	0.00078	0.00208
1.8		0.00001	0		0.01148	0.00105	0.00193
2.0	⋮	0			0.00444	0.00123	0.00020
2.2					0.00635	0.00108	0.00173
2.4		⋮	⋮	⋮	0.00521	0.00066	0.00202
2.6					0.00205	0.00019	0.00073
2.8					0.00407	0.00021	0.00084
3.0	0	0	0	0	0.00039	0.00051	0.00119

TABLE IX

t	$x_3(t)$	δ_{AV}^{32}	δ_{AV}^{64}	δ_{AV}^{128}	$\delta_{S_3}^{4}$	$\delta_{S_3}^{8}$	$\delta_{S_3}^{16}$
0.2	1.0	0	0	0	0.01131	0.00517	0.00235
0.4		0	0	0	0.01278	0.00471	0.00022
0.6	⋮	0.0034	0	0	0.00628	0.00373	0.00202
0.8		0.01658	0.00386	0.00039	0.00472	0.00100	0.00164
1.0	1.0	0.12652	0.09231	0.06676	0.04871	0.03399	0.01807
1.2	1.36	0.00115	0.00751	0.00665	0.02598	0.00193	0.00149
1.4	1.64	0.02780	0.01544	0.00781	0.02266	0.00337	0.00228
1.6	1.84	0.03040	0.01559	0.00781	0.00015	0.00168	0.00248
1.8	1.96	0.02546	0.01424	0.00760	0.01932	0.00086	0.00153
2.0	2.0	0.00207	0.00073	0.00026	0.01007	0.00266	0.00049
2.2		0.02246	0.01347	0.00746	0.00901	0.00349	0.00223
2.4	⋮	0.02950	0.01547	0.00781	0.01127	0.00375	0.00198
2.6		0.03098	0.01562	⋮	0.00071	0.00321	0.00068
2.8		0.03122	0.01562		0.00714	0.00211	0.00049
3.0	2.0	0.03125	0.01562	0.00781	0.00465	0.00076	0.00059

$t < 1 - \epsilon$ and $t > 1 + \epsilon$, ϵ some positive number, is due to the specific choice of the initial data.

EXAMPLE 5. For the scalar equation

$$\dot{x}(t) = 5x(t) + x(t - 1)$$

with initial data

$$x(\theta) = 5 \qquad \text{for} \quad \theta \in [-1, 0]$$

we give the numerical data for AV, S_1, and S_3 in Tables X and XI. The true solution is

$$x(t) = 6e^{5t} - 1 \qquad \text{for} \quad t \in [0, 1]$$

and

$$x(t) = \{x(1) - \tfrac{1}{5} + 6(t-1)\}\, e^{5(t-1)} + \tfrac{1}{5}$$

for $t \in [1, 2]$. The initial data are in Z^N, $N = 1, 2, \ldots$, for all three approximations AV, S_1, S_3. The numerical results show that the approximation by AV is initially better than by S_1 and S_3. This appears to be the case in all examples where the initial data are in the subspaces Z^N for AV, S_1, and S_3. But the relative error increases much faster with increasing time for AV than for S_1 and S_3. For instance, for $N = 32$, S_1 will do better than AV shortly after $t = 2$.

TABLE X

t	$x(t)$	δ^8_{AV}	δ^{16}_{AV}	δ^{32}_{AV}	$\delta^8_{S_1}$	$\delta^{16}_{S_1}$	$\delta^{32}_{S_1}$
0	5.0	0	0	0	0	0	0
.25	19.9421	0.0001	0.0001	0.0001	0.2853	0.0756	0.0192
.50	72.0950	0.0160	0.0007	0.0001	1.5512	0.4090	0.1037
.75	254.126	0.428	0.081	0.011	7.352	1.937	0.492
1.00	889.479	4.369	1.481	0.539	32.380	8.531	2.163
1.25	3109.32	26.65	10.46	4.47	136.14	35.93	9.11
1.50	10870.4	130.9	53.3	23.4	555.8	147.0	37.3
1.75	38004.6	589.2	244.6	108.7	2219.9	588.3	149.4
2.00	132871.	2519.	1058.	474.	8721.	2316.	588.

TABLE XI

t	$x(t)$	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	$\delta^{32}_{S_3}$	$\delta^{64}_{S_3}$
0.2	15.309691	0.030796	0.004862	0.000289	0.000450	0.000032
0.4	43.334337	0.051843	0.003990	0.001761	0.000128	0.000303
0.6	119.513222	0.260592	0.012450	0.001268	0.000888	0.001223
0.8	326.588900	0.781997	0.018429	0.000120	0.004050	0.003180
1.0	889.478955	2.388011	0.094175	0.002117	0.013687	0.008994
1.2	2420.772761	7.268330	0.243622	0.011288	0.038706	0.076608
1.4	6588.865818	21.936935	0.709869	0.028940	0.155481	0.296237
1.6	17934.15321	65.33753	2.00746	0.07668	0.54728	1.26327
1.8	48815.25691	193.16251	5.69496	0.52984	1.80974	4.94164
2.0	132871.3779	566.4524	16.8275	1.4527	4.9407	17.0516

The numerical data also indicate accumulation of errors for S_3 if N and t are large. The rate of convergence for S_1 again is like $1/N^2$. For S_3 it seems to be like $1/N^{4+\epsilon}$, $\epsilon > 0$.

From the numerical results presented here we can draw the following conclusions:

(a) The rate of convergence for S_1 in all cases was like $1/N^2$ compared with $1/N$ or less for AV. This superiority is not only true for examples where we considered just an initial value problem but also for examples (not presented in this paper) where optimal control problems and identification problems were considered. For S_3 conclusions to be made from the data are not as clear. In the case of the scalar equation of Example 5 the rate of convergence seems to be better than $1/N^4$. For the three-dimensional system of Example 4, $N = 16$ gives just a slightly better approximation than $N = 4$. A possible explanation for this behavior is the Eq. (3.9) must be dealt with more carefully than we did in our preliminary calculations.

(b) For the same value of N, the accuracy of approximation over large t-intervals for S_1 and S_3 is considerably better than for AV. The latter scheme does better initially if the initial data are already in the subspaces Z^N for AV and in some exceptional cases where the simple algebraic structure of AV is advantageous (see Example 4).

(c) For S_1 implementation of the algorithms on a computer is almost as easy for AV. For S_3 the matrices appearing in the algorithm are more complicated. On the other hand in many cases the results using AV were comparable in accuracy to those obtained using S_3 with $N = 4$ only if $N > 100$ was taken in the AV approximation.

REFERENCES

1. H. T. BANKS, Approximation of nonlinear functional differential equation control systems, *J. Optimization Theory Appl.*, in press.
2. H. T. BANKS AND J. A. BURNS, An abstract framework for approximate solutions to optimal control problems governed by hereditary systems, *in* "International Conference on Differential Equations" (H. Antosiewicz, Ed.), Academic Press, New York, 1975, pp. 10–25.
3. H. T. BANKS AND J. A. BURNS, Hereditary control problems: Numerical methods based on averaging approximations, *SIAM J. Control Optimization* **16** (1978), 169–208.
4. H. T. BANKS AND J. A. BURNS, Approximation techniques for control systems with delays, *in* "Proceedings Conference on Methods of Math. Programming, Zakopane, Poland, September, 1977," in press.
5. V. BARBU, "Nonlinear Semigroups and Differential Equations in Banach Spaces," Noordhoff, Leyden, 1976.

6. J. G. Borisovič and A. S. Turbabin, On the Cauchy problem for linear nonhomogeneous differential equations with retarded arguments, *Soviet Math. Dokl.* **10** (1969), 401–405.
7. J. A. Burns and E. M. Cliff, Methods for approximating solutions to linear hereditary quadratic optimal control problems, *IEEE Trans. Automatic Control* **23** (1978), 21–36.
8. E. M. Cliff and J. A. Burns, Parameter identification for linear hereditary systems via an approximation technique, *in* "Proceedings, Workshop on the Linkage between Applied Mathematics and Industry, March 1978," in press.
9. F. Kappel and W. Schappacher, Autonomous nonlinear functional differential equations and averaging approximations, *J. Nonlinear Analysis: Theory, Methods Appl.* **2** (1978), 391–422.
10. A. Pazy, "Semigroups of Linear Operators and Applications to Partial Differential Equations," Math. Dept. Lecture Notes, Vol. 10, University of Maryland, College Park, Md., 1974.
11. D. Reber, "Approximation and Optimal Control of Linear Hereditary Systems," Ph. D. Thesis, Brown University, Providence, R. I., November 1977.
12. D. Reber, A finite difference technique for solving optimization problems governed by linear functional differential equations, *J. Differential Equations* **32** (1979), 193–232.
13. P. Rotter, "Splineapproximation von Differenzen-Differentialgleichungen," Diplomarbeit, Technische Universität Graz, Graz, Austria, June 1978.
14. M. H. Schultz, "Spline Analysis," Prentice–Hall, Englewood Cliffs, N. J., 1973.
15. G. F. Webb, Functional differential equations and nonlinear semigroups in L^p-spaces, *J. Differential Equations* **20** (1976), 71–89.
16. K. Yosida, "Functional Analysis," Springer–Verlag, New York, 1974.